AFRL-ML-WP-TP-2006-450

FUSING COMPETING PREDICTION ALGORITHMS FOR PROGNOSTICS (PREPRINT)

Kai Goebel, Neil Eklund, and Pierino Bonanni

MARCH 2006

Approved for public release; distribution is unlimited.

STINFO COPY

This work was funded in whole or in part by Department of the Air Force contract HR0011-04-C-0001. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the U.S. Government.

MATERIALS AND MANUFACTURING DIRECTORATE
AIR FORCE RESEARCH LABORATORY
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7750

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS**.

1. REPORT DATE ***(DD-MM-YY)***	**2. REPORT TYPE**	**3. DATES COVERED** ***(From - To)***
March 2006	Conference Paper Preprint	

4. TITLE AND SUBTITLE FUSING COMPETING PREDICTION ALGORITHMS FOR PROGNOSTICS (PREPRINT)	**5a. CONTRACT NUMBER** HR0011-04-C-0001 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER** N/A
6. AUTHOR(S) Kai Goebel, Neil Eklund, and Pierino Bonanni	**5d. PROJECT NUMBER** M02R **5e. TASK NUMBER** 30 **5f. WORK UNIT NUMBER** 00
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) GE Global Research One Research Circle Niskayuna, NY 12309	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Materials and Manufacturing Directorate Air Force Research Laboratory Air Force Materiel Command Wright-Patterson AFB, OH 45433-7750	**10. SPONSORING/MONITORING AGENCY ACRONYM(S)** AFRL-ML-WP **11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)** AFRL-ML-WP-TP-2006-450

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
This work was funded in whole or in part by Department of the Air Force contract HR0011-04-C-0001. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the U.S. Government.
This work has been submitted for publication in the Proceedings of the 2006 IEEE Aerospace Conference, published by IEEE. PAO Case Number: AFRL/WS 06-0085, 09 Jan 2006.

14. ABSTRACT
Two fundamentally different approaches can be employed to estimate remaining life in faulted components. One is to model from first principles the physics of fault initiation and propagation. Such a model must include detailed knowledge of material properties, thermodynamic and mechanical response to loading, and the mechanisms for damage creation and growth. Alternatively, an empirical model of condition-based fault propagation rate can be developed using data from experiments in which the conditions are controlled or otherwise known and the component damage level is carefully measured. These two approaches have competing advantages and disadvantages. However, fusing the results of the two approaches produces a result that is more robust than either approach alone. In this paper, we introduce an approach to fuse competing prediction algorithms for prognostics. Results presented are derived from rig test data wherein multiple bearings were first seeded with small defects, then exposed to a variety of speed and load conditions similar to those encountered in aircraft engines, and run until the ensuing material liberation accumulated to a predetermined damage threshold or cage failure, whichever occurred first.

15. SUBJECT TERMS
Bearing damage, prognostics, diagnostics, Prognostic Reasoner, life prediction, fault propagation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Andrew H. Rosenberger
Unclassified	Unclassified	Unclassified	SAR	16	**19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Fusing Competing Prediction Algorithms for Prognostics

Kai Goebel, Neil Eklund, Pierino Bonanni
GE Global Research
One Research Circle
Niskayuna, NY 12309
518-387-4194
{goebelk, eklund, bonanni}@research.ge.com

Abstract— Two fundamentally different approaches can be employed to estimate remaining life in faulted components. One is to model from first principles the physics of fault initiation and propagation. Such a model must include detailed knowledge of material properties, thermodynamic and mechanical response to loading, and the mechanisms for damage creation and growth. Alternatively, an empirical model of condition-based fault propagation rate can be developed using data from experiments in which the conditions are controlled or otherwise known and the component damage level is carefully measured. These two approaches have competing advantages and disadvantages. However, fusing the results of the two approaches produces a result that is more robust than either approach alone. In this paper, we introduce an approach to fuse competing prediction algorithms for prognostics. Results presented are derived from rig test data wherein multiple bearings were first seeded with small defects, then exposed to a variety of speed and load conditions similar to those encountered in aircraft engines, and run until the ensuing material liberation accumulated to a predetermined damage threshold or cage failure, whichever occurred first.

Table of Contents

1. Introduction 1
2. Background 2
3. Integrated Reasoner Operating Modes 3
4. Models 5
5. Prognostic Reasoner Module 5
5. Results and Discussion 7
6. Summary & Conclusions 8
Acknowledgments 8
References 8
Biography 9

1. Introduction

Prognostics plays a pivotal role in integrated systems health management (ISHM). Estimating remaining component life with uncertainty bounds that are narrow enough to allow system operation after a fault has been detected offers the prospect for increased system safety along with more cost effective maintenance strategies. The latter include performing on-demand maintenance, a departure from the traditional, actuarial-based practice in which components are managed to life limits based upon fleet wide statistics and average expected usage. The traditional approach is necessarily conservative, requiring the replacement of parts irrespective of how much of their useful life is actually expended. In contrast, a condition-based parts replacement strategy [Orsagh et al., 2003] results in reduced cost of ownership with the same safety margin. The whole paradigm of fleet management could be changed because it would be possible to not only perform maintenance at a convenient place and time, taking into account variables such as part and staff availability, shop loading, and other factors, but also to plan future missions more reliably. To that end, a DARPA-sponsored program -- of which the work reported here is a part -- addresses engine prognosis using advanced physics-based models, state-awareness sensors, and a prognostic reasoner to compute component capability, to quantify prediction-related variability, and to provide system-wide capability assessment [Littles and Buczek, 2004].

Remaining life prediction is at the core of prognostics. However, there are considerable imponderables of such a life prediction. First and foremost, these result from the uncertainties surrounding the future use of the component of interest. A component has generally shorter remaining life when it is subjected to higher load conditions than when it experiences lesser load conditions. However, unless the system is run under constant load conditions or goes through repetitive load cycles, the detailed nature of future load conditions will have an impact on the quality of the remaining life estimate. There are numerous other conditions that may influence the remaining life, depending on the component and fault mode. For bearings, these may include speed, temperature, humidity, external vibration, contamination, lubrication, and other factors.

The physics-based prognostic models deal with mechanisms

[1] 0-7803-8870-4/05/$20.00© 2005 IEEE
[2] IEEEAC paper #593, Version 1, Updated October 12, 2005

governing incipient damage at the material level, factoring in both full finite-element and reduced-order formats. The state awareness sensors measure material and system damage state, identify engine operation conditions, and update model predictions with advanced signal-acquisition and signal-conditioning methodologies. Data from rig tests were used to develop in parallel an empirical model of spall growth rate (as a function of speed and load). Finally, the prognostic reasoner fuses sensor and model-based information to assess residual component capability, calculate the uncertainty level for system predictions, and project a safe operational envelope for near-term engine usage [Littles and Buczek, 2004].

In an earlier paper [Goebel et al., 2005] we presented an architecture for the prognostic reasoner demonstrated on a bearing system (Figure 1). This paper will focus on the prognostic reasoner and specifically on one way to fuse competing remaining life estimates. Generally, the reasoner is represented as a multi-layered architecture comprising pre-processing, analysis, and post-processing steps. These steps are partitioned into modules where each module performs supporting tasks for both information processing and uncertainty management. The focus of this paper is on a method within the analysis module.

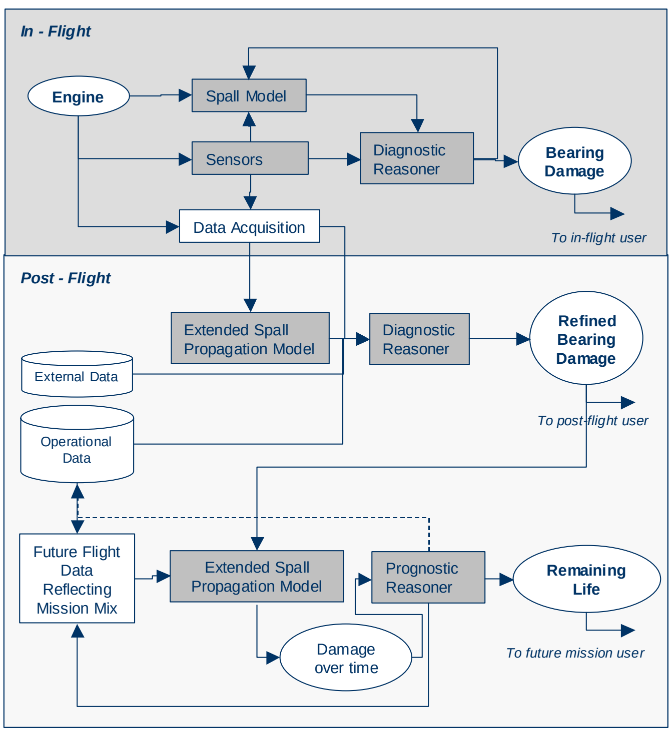


Figure 1 - Interactions of Integrated Bearing Reasoner Modules

2. BACKGROUND

Bearing Damage

During bearings operation, initially localized spalls can initiate that may grow and ultimately result in loss of function. Important factors affecting damage initiation and damage propagation are changes in bearing loads, speeds, and environment. Lubrication, presence of material defects, surface degradation, and external contamination all factor in to the bearing environment. Subsurface fatigue cracks are induced at locations of peak shear stress, become surface-connected, and lead to eventual liberation of material. It is important to assess the microstructural evolution, environmental embrittlement, cyclic hardening, and residual stress to calculate the propagation of bearing damage. The current state is determined by feeding direct sensor data and indirect parameters computed from sensor data into an ensemble of diagnostic algorithms as a basis for input to the fault-evolution and life models [Littles and Buczek, 2004]. The algorithms arrive at their conclusion either by direct measurement, models supported by measurements, or are simply triggered by measurements. The information sources that the reasoner relies on may be updated at different intervals during or between flights and may have different prediction horizons.

Sensors for Bearing Prognostics

Prognostics is about the estimation of remaining useful life under particular assumptions of future use. Sensor measurements provide instantaneous feedback on current damage levels and form the foundation for prognostic estimates. Ideally, features derived from sensor measurements would have monotonically changing properties that accurately reflect increasing component damage and be provided irrespective of external conditions. However, in practice this is nearly never the case: features reflect the noise inherent in sensed data and react differently during particular stages of damage evolution (e.g., some are useful for fault detection, but not for damage growth tracking).

Oil debris monitor features, such as particle counts, have excellent tracking properties that are invariant to changes of environmental parameters [Dempsey et al., 2002]. However, they may be not as suitable to identify fault initiation because their resolution is too low for small damage levels. Better sensors for fault initiation and initial fault growth tracking may be vibration sensors that have the promise to pick up smaller damage levels. Features from various transforms such as Fourier, Hilbert, and Wavelets can be useful in detecting and categorizing incipient faults. The vibration sensor’s capacity for early detection comes at the price of sensitivity to environmental effects [Dempsey et al., 2002] that are sometimes difficult to quantify or correct. In an aircraft engine, and in particular under conditions of military use, these changes can be significant.

It is thus expedient to aggregate vibration and oil debris information to take advantage of the benefits of both. The fusion of information from oil debris and vibration information, along with knowledge about system and machinery history can result in interactions that may improve the confidence about system condition [Byington et al., 1999]. Howard and Reintjes [Howard and Reintjes, 1999] describe the benefits of using several information sources for fault detection, and discuss oil debris and vibration for helicopter gearboxes in particular. Byington et al. [Byington et al., 1999] describe a fusion technique that correlates the failure mode phenomena with appropriate features. Dempsey et al. [Dempsey et al., 2002] report on the use of fuzzy logic to integrate oil debris and vibration information for gearbox faults where the output was quasi-action recommendations such as “OK, inspect, shutdown”.

Diagnostics

Prognostics is reliant on diagnostics to provide a trigger point. That is, no prognostic estimates are calculated before diagnostics has detected a fault condition. In the absence of abnormal conditions – or fault conditions – the best estimates for remaining component life are often fleet wide statistics expressed by Weibull curves or other suitable mechanism. Condition-based systems depend on reliable fault diagnostics to initiate the prognostic algorithms.

Prognostics

The remaining useful life (RUL) estimates are in units of time until the likelihood of failure reaches a particular threshold. RUL is often estimated indirectly via the calculation of a metric that, when exceeding a particular threshold, indicates imminent component failure. In the context of bearing race spall, this metric could be spall length. When spall length surpasses the ball spacing, damage accumulates rapidly; bearing cage failure occurs soon after this threshold has been breached.

Two fundamentally different approaches can be employed to estimate future damage. One is to model from first principles the physics of the system as well as the fault propagation for given load and speed conditions. Such a model must include detailed knowledge of material properties, thermodynamic behavior, etc. Alternatively, an empirical experience-based model can be employed wherein data from experiments at known conditions and component damage level are used to build a model for fault propagation rate. Such a model relies heavily on a reasonably large set of experiments that sufficiently explores the load and speed space.

Prognostic Information Fusion

The two approaches for estimating future damage have various advantages and disadvantages. The physics-based model relies on the assumption that the fault mode modeled using the specific geometry, material properties, temperature, load, and speed conditions will be similar to the actual fault mode. Deviation in any of those parameters will likely result in an error that is amplified over time. In contrast, the experience-based model assumes that the data available sufficiently maps the space and that interpolations (and extrapolations) from that map can capture the fault rate properly. It can be beneficial to fuse the output of both methods to produce a more robust and more accurate result. Finding synergy in using different information sources to assess system states has a long tradition within the fields of multivariate statistics and pattern recognition.

In addition to fusing a damage estimate, the associated uncertainty needs to be aggregated as well. This is a critical task because the resulting estimate needs to be within uncertainty bounds that allow for decision making at a desired risk level. If the uncertainty bounds are very wide, the resulting time-of-failure estimate at the acceptable risk level may be too early to provide any benefit to the decisioning process. That is, there would be no advantage of prognostics compared to a reactionary diagnostics system alone. Uncertainty bounds ideally are tight but need to reflect true output variability.

Prognostic Fusion Techniques

The aggregation of future damage estimates is not just a question of averaging the various values. Rather, the fusion method should be able to incorporate a number of different measures that inform about the reliability of the estimate, their expected accuracy, and various other uncertainty measures. These measures in turn may be a function of different variables such as time, where in the load/speed space the estimate is performed, known shortcomings or strength in some areas of that space, etc. In the example described by Orsagh et al. [Orsagh et al., 2003], performance improvement is accomplished when weights for the information sources are dynamically allocated depending on whether the component is considered early or late in its remaining useful life cycle. Garga et al. [Garga et al., 2001] describe a hybrid reasoning approach that integrates domain knowledge with test and operational data from an industrial gearbox. There, domain knowledge is expressed as a rule-base, and then used to train a feedforward neural network.

3. INTEGRATED REASONER OPERATING MODES

As mentioned above, the prognostic reasoner considered here is really a set of reasoners that will operate at various times during and after the flight. Depending on the time during or after a mission, its tasks will vary from aggregation of damage information to supporting the calculation of a remaining life estimate.

In-Flight and Post-Flight Diagnostic Modes

During the flight, there are a number of features derived from sensors that inform about the presence of bearing damage. Specifically, this information encompasses features derived from accelerometers that measure and assess vibration. Furthermore, information from debris monitoring devices is also used as a sensor-based input to the reasoner. In addition, a spall propagation model will provide information about the size and rate of increase of spalls. This model will use triggers from the reasoner to initiate its operation. That is, it will be dormant in the absence of evidence of bearing damage, and fleet-wide statistics on bearing fatigue are used for low-level damage accumulation. Therefore, the overall reasoner will initially have to reliably provide diagnostic information about bearing damage to the spall propagation model. Once bearing damage has been established and the spall propagation model has been triggered, it will also need to integrate the information from the spall propagation model with the vibration and debris information. The output of the in-flight diagnostic reasoner is a damage estimate. The post-flight diagnostic reasoner performs similar functions with the difference that it is not encumbered by computational constraints. That is, it will run full-order models that provide a more refined damage estimate.

Prognostic Mode

The prognostic models can be run either on-board or on-ground, depending on whether there is a need for short term outlook (in which case the prognostic reasoner would be executed on-board) or whether there is a need for a longer-term outlook (in which case it makes more sense to run the prognostic reasoner on-ground). If a fault has been detected, the prognostic functions are executed on a set of future missions. Specifically, missions characterized in part by sequences of load, speed, and ambient conditions are used as input to the physics-based spall propagation model. In conjunction with the current damage state, the output of the spall propagation model will provide a damage profile into the future. In parallel, the future mission profile is also used to execute the experience-based model, which gives its independent future damage estimate.

In addition to the damage estimate, each model is assigned a quality assessment that can be interpreted as confidence. These confidences are computed based on *a priori* performance of the models. That is, the models may be known to have a different performance within different regions of the load-speed mission space. Additionally, the models may be known to produce biases at different damage levels or at different damage rate levels. Furthermore, the further out into the future the prediction is being made, the less likely it is correct. While confidence intervals may capture the possible variability, the quality assessment captures other sources of uncertainty. For example, the top set of axes of Figure 2 shows some data fit with a linear model. The dashed lines show the 95% confidence interval *of the model*. However, we can plainly see (second set of axes) in this simple case that a linear fit is a poor approximation of the data (it is less obvious in space of higher dimensionality). If this model was for example driven by first principles to this particular form, the confidence interval of the model alone does not do us much good. However, if we are able to take into account the *quality* of the model (e.g., derived by examining performance of the model) for particular regions of the search space (or other factors, e.g., time), we can arrive at a much better fusion of the data.

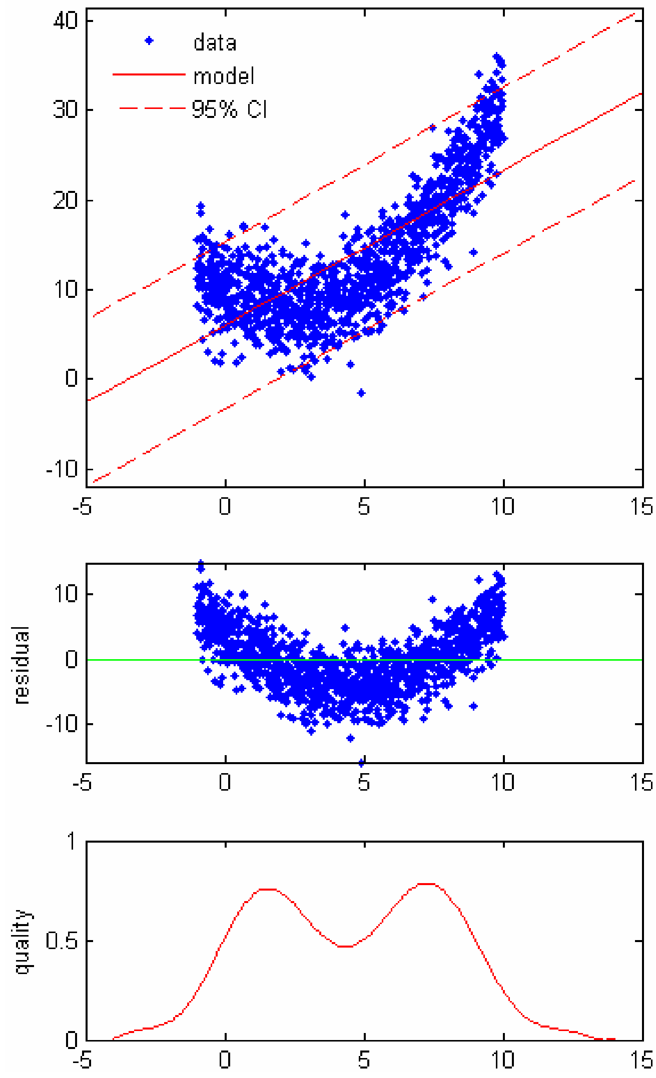


Figure 2 – Data with linear fit, residual, and derived quality assessment

The modeled damage over time and the quality assessment over time from each model are then forwarded to the aggregation module. Figure 3 illustrates the operation of the prognostic reasoner. Fundamentally, the prognostic reasoner supervises the execution of the different prognostic models, makes corrections where desired, and assigns a quality assessment. It then aggregates the different estimates. There are different ways in which the reasoner can operate based

on user demand. In one instantiation, it will report both the profile of remaining life and information on whether the envisioned missions can be completed without exceeding the acceptable damage limit. In another instantiation, it will provide information back to the mission generation process to prompt for additional mission runs when damage limits have not been reached. The goal of executing the damage propagation model with additional runs is to determine the damage propagation profile and to find the remaining life limit.

As mentioned before, if no fault has been detected, the prognostic module is bypassed and is replaced by fleet statistics that are compiled on bearing fatigue data.

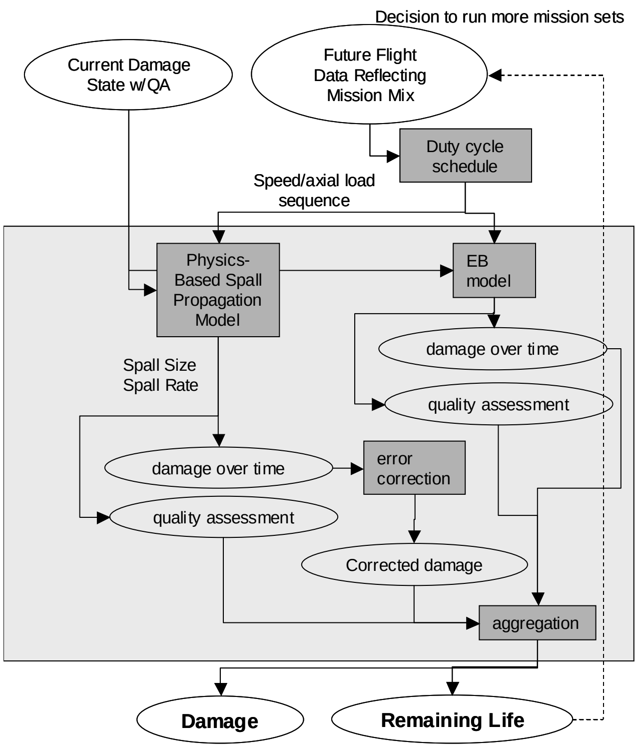


Figure 3 - Prognostic Reasoner

4. MODELS

Two models are fused in the prognostic reasoner, a physics-based (PB) model and an experience-based (EB) model.

The PB model for the initiation and propagation of bearing fatigue spalls uses historic and estimated future operating conditions to determine future bearing condition and returns a probability density function of the bearing remaining useful life. This model is based on first principles approaches such as damage mechanics to track material microstructure changes and eventual loss during the spall propagation phase. It takes into account material properties, bearing geometry, surface interaction, lubrication, and variable operating conditions.

The EB model is an empirical fit of data from seven experiments at five points in the speed and load space. Spall length is calculated:

$$l_{spall} = 10^{\log 10(l_{spall_{t=0}}) + \sum_{t=0:dt:current} rate(t)*dt}$$

where

$$rate = 10^{f(speed(t), load(t))}.$$

Spall growth rate is exponential, with rate an empirical function of speed and load (Figure 4).

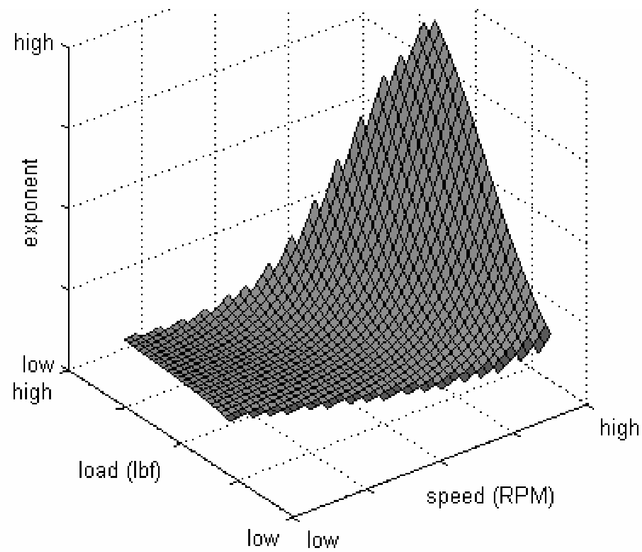


Figure 4 - Response surface of experience-based model.

5. PROGNOSTIC REASONER MODULE

The primary goal of the prognostic reasoner is to negotiate the different damage estimates and to decide whether another set of mission parameters needs to be executed for another damage estimate further in the future. The aggregation of different estimates is the focus of this paper.

There are numerous approaches such as bagging and boosting [Freund and Schapire, 1999], Dempster-Shafer [Smets, 1994], model-based approaches [Nelson and Mason, 1999], fuzzy fusion [Loskiewicz and Uhrig, 1994] or statistics-based approaches [Rao, 2000] that attempt to address the core aggregation functions. However, it has to be realized that the aggregation itself is only one function of the overall reasoner. In addition to combining information, it has to be ensured that the information that is being used provides the maximum information content. There are a number of issues that need to be dealt with prior to the actual aggregation. Specifically, the information needs to be checked for consistency, and it needs to be cleaned of outliers, noise, faulty or otherwise bad sensor information, it needs to be conditioned and formatted to allow a proper

comparison. In addition to that, special cases need to be taken into account that, depending on the situation, should be done either before or after the actual aggregation step. To assist in these tasks, we suggest employing a sequential and parallel multi-layered configurations strategy. Elements from this configuration strategy have been proven successful in diagnostic fusion environments within project IMATE [Ashby and Scheuren, 2000]. There, a hierarchical, multi-layer architecture [Goebel, 2001] was demonstrated that implemented some of these concepts. Information from various diagnostic models and evidential information sources was combined and manipulated through a series of steps that increased and decreased the weight given to the information sources according to the strategies implemented in the respective layers of the fusion process.

In the following section we describe algorithmic concepts of the in-flight prognostic reasoner. In contrast to a diagnostic reasoner that has the task to determine the presence of a fault and therefore has as its output the fault category and perhaps an associated confidence, the in-flight prognostic reasoner needs to assess the presence of an initial fault condition and to report on the overall damage level plus an associated uncertainty. The most fundamental difference is in the second task, namely producing a damage assessment output that is in continuous format. This means that different aggregation techniques will need to be employed.

Fusion is performed in the analysis module. We have tested a number of different fusion techniques including weighted averaging and adaptive neuro-fuzzy inference systems. The latter has the advantage of automated learning capability, while the former relies on the user to provide the appropriate weights. The two approaches mentioned both arrived at satisfactory results. Ultimately, however, the final fusion strategy needs to also include a provision to aggregate different measures of uncertainty. For this purpose, a Dempster-Shafer-based regression method [Petit-Renaud and Denoeux, 2004] was used. The basic concept of regression is to determine a functional relationship between two or more correlated variables that is often empirically derived from data and is used especially to predict values of one variable when given values of the others; specifically, a function that yields the mean value of a random variable under the condition that one or more independent variables have specified values.

In non-parametric regression (NPR), no assumptions about the underlying functional form are made. NPR is characterized by low bias (i.e, it can easily represent underlying function) but at the expense of high variance (i.e., the model will change from realization to realization of the data). That in turn may change the response dramatically depending on data. The simplest idea is the k-nearest neighbor regression that results in good fit, but huge variance and discontinuous behavior. Kernel regression overcomes some of these shortcomings by locally weighting members closer to the value in question. The operative equation of Nadaraya-Watson [Watson, 1964; Nadaraya, 1964] kernel regression is

$$\hat{f}(x) = \frac{\sum_{i=1}^{N} K_\lambda(x_0, x_i) y_i}{\sum_{i=1}^{N} K_\lambda(x_0, x_i)}$$

Classical regression techniques (kernel, MLP, RBF, splines, linear, etc.) assume perfect knowledge of y (both precise and certain). However, these techniques do not work optimally if knowledge of sensor measurement y is imprecise due to limited precision and accuracy of sensors, and if sensor measurement y is uncertain (e.g., due to sensor failure). The issue is exacerbated when there are multiple sensors with different sensitivities and reliabilities. In situations where the probe point is very different from that employed in the training set it might be desirable to have mechanisms to cast doubt on the validity of the output.

Dempster Shafer regression [Petit-Renaud and Denoeux, 2004] (DSR) provides a prediction of the output in form of a fuzzy belief assignment. This assignment is defined as a collection of fuzzy sets of values with associated masses of belief. The output is computed using a nonparametric, instance-based approach: evidence samples $e_i = (x_i, m_i)$ in the neighborhood of the input vector x are sources of partial information on the response variable. The evidence samples can be represented by a fuzzy belief assignment $m_y[x, e_i]$. Relevance of the evidence with respect to y is assumed to be dependent on the dissimilarity to y. If x is "close" to x_i according to a given metric $\|.\|$, y is expected to be close to y_i, which makes example e_i quite relevant to predict the value of y. On the contrary, if x and x_i are very dissimilar, example e_i provides only marginal information regarding the value of y. Therefore, neighborhood evidence input elements are discounted as a function of their distance to x. They are then pooled using Dempster's rule of combination. While the method can cope with heterogeneous training data, the more important characteristics in this context is the formalism for modeling both unreliable and imprecise information provided by multi-sensor systems.

DSR determines the value of sensor measurement y at a given time by discounting the belief mass of each observation by:

$$\phi\left(\left|x - x_i\right|\right) = \gamma e^{-\frac{(x-x_i)^2}{\Theta^2}}$$

where:

γ is a tuning parameter (usually >= 0.9)

Θ is a scale parameter, commonly set using cross validation on training data

Next, the discounted belief masses are combined using appropriate version of DS combination. When there are many data points, the computational overhead can become considerable. A remedy is using only the k-nearest

neighbors to reduce the complexity of the calculation with little loss of accuracy.

The future estimated variability of the estimators was folded into the Dempster-Shafer model by adding the anticipated variability to the output.

5. Results and Discussion

The prognostic reasoner has been tested on a sequential set of experiments that model a simulated, cyclic mission profile. Figure 5 shows the assembled load and speed history, which was reflective of about 40 cycles in the load-speed space, with dwells at certain setpoints. An indent was added to the outer race of a production bearing, which was then run under those conditions. The bearing was examined several times during the course of the test, and actual spall length was recorded. The test ran to cage failure.

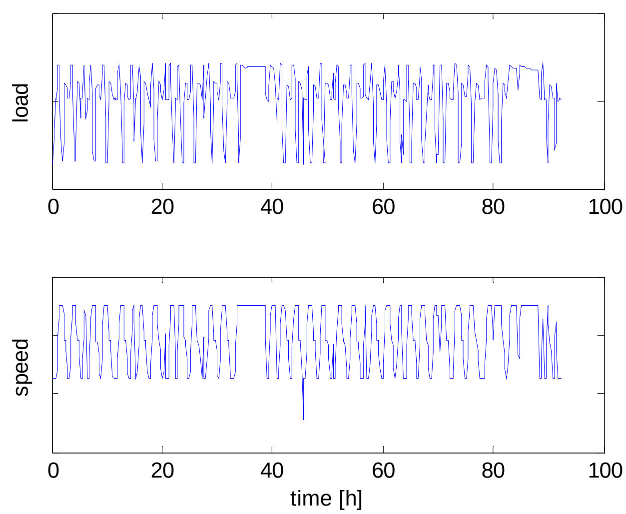


Figure 5 – Test Profile (load and speed)

Figure 6 shows the subjective confidences that were assigned to the estimators in the forward mode. The forward mode can be executed at any time. For illustrative purposes, we chose t=70 hours as the starting point. The fundamental characteristic of the forward confidences is that they drop as a function of time. In addition, there is an *a priori* bias assigned to the different confidences, which in turn reflects the accuracy of the models as observed during testing. Figure 8 shows the output of the prognostic reasoner run forward to t = 93 hours. The green line reflects the output of the experience-based model in forward mode. The blue line is the output of the physics-based model in forward mode. The stars are the measurements taken during the experiment. Although they would not be available during an actual forward mode, they are shown here to illustrate agreement with real damage. Both the experience-based model and the physics-based model show an increase of damage over time. At smaller time scales (not shown here), one can see how the different operating conditions have varied impact on the forward model. The physics-based model has a larger bias that leads to underestimation of the actual damage.

Also shown is the reasoner output. Specifically, the solid black line is the 50th percentile of the reasoner. The dashed black lines are the 5th and 95th percentiles of the reasoner. The 95th percentile line crosses the critical damage level of 7% at about 75 hours. Depending on the risk tolerance, this can be used by operators to either schedule maintenance or to alter the planned mission sequence -- through correction of the load/speed profile -- to achieve a longer time to failure.

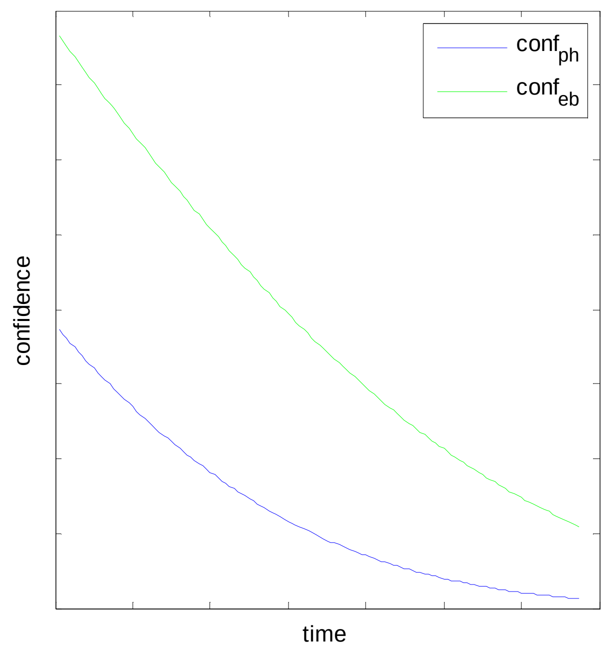


Figure 6 – Subjective confidence for competing estimates

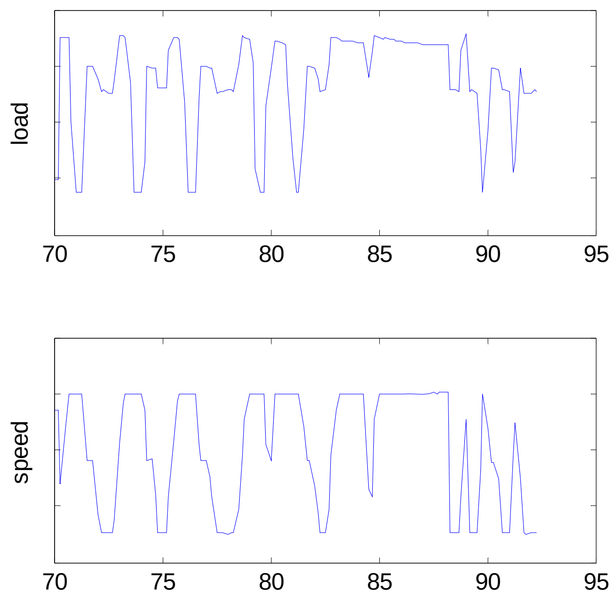


Figure 7 – Anticipated Future Load Profile

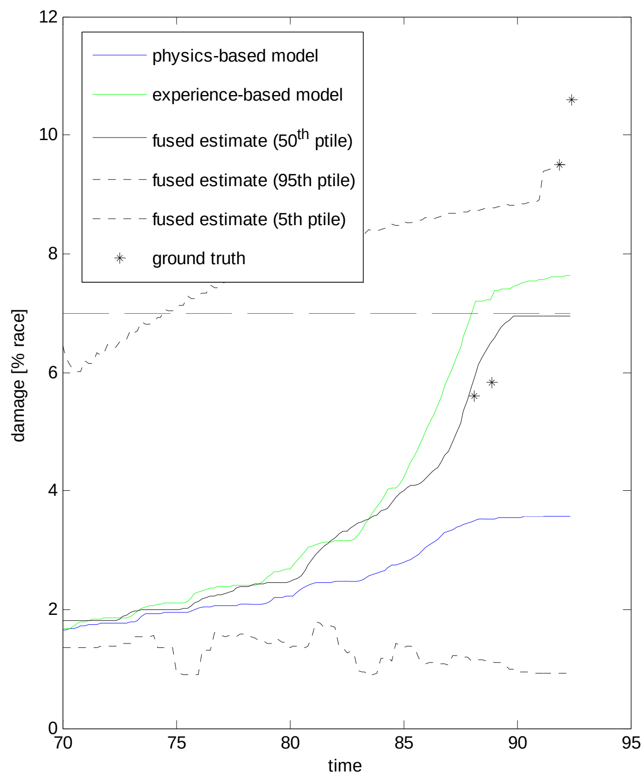


Figure 8 - Prognostic Output at t = 70 hours

6. SUMMARY & CONCLUSIONS

This paper describes how two fundamentally different methods can be employed to estimate remaining life and how their independent estimates can be fused. One method uses first principles to model fault propagation through consideration of the physics of the system. The other method is an empirical model using data from experiments at known conditions and component damage level to estimate condition-based fault propagation rate. These two approaches are fused to produce a result that is more accurate and more robust than either method alone. The fusion method employs a Dempster-Shafer regression that – in addition to the damage estimates – takes advantage of subjective quality assessments that quantify the uncertainty of the estimates at any time. We present results from rig tests where a bearing was run under mission typical flight profiles. Spall was initiated and bearing spall growth was carefully monitored. Results from these tests were compared to the prognostic estimates of the reasoner.

ACKNOWLEDGMENTS

This Material is based upon work supported by the Defense Advanced Research Projects Agency, Defense Sciences Office (DSO), Engine System Prognosis, ARPA Order No. Q517/00, Issued by DARPA/CMO under Contract Nos. HR0011-04-C-0001. The views, opinions, and/or findings contained in this article are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the US Air Force, or the Department of Defense.

REFERENCES

M. Ashby and W. Scheuren, “Intelligent Maintenance Advisor for Turbine Engines (IMATE)”, Proceedings of the IEEE Aerospace Conference, 11.0309, 2000.

C. Byington, T. Merdes, and J. Kozlowski, “Fusion Techniques for Vibration and Oil Debris/Quality in Gearbox Failure Testing”, Proceedings of the International Conference on Condition Monitoring ’99, pp. 113 – 118, 1999.

P. Dempsey, R. Handschuh, and A. Afjeh, “Spiral Bevel Gear Damage Detection Using Decision Fusion Analysis”, Proceedings of the Fifth International Conference on Information Fusion, Fusion 2002, vol. 1, pp. 94-100, 2002.

Y. Freund and R. Schapire, “A Short Introduction to Boosting”, J. Japanese Society Artificial Intelligence, Vol. 14, No.5, pp. 771-780, 1999.

A. Garga, K. McClintic, R. Campbell, C.-C. Yang; M. Lebold, T. Hay, C. Byington, “Hybrid Reasoning for

Prognostic Learning in CBM Systems”, IEEE Proceedings Aerospace Conference, vol. 6, 10-17 March 2001, pp. 2957 – 2969, 2001.

K. Goebel, “Architecture and Design of a Diagnostic Information Fusion Tool”, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 15 (4), Sept. 2001, pp. 335-348, 2001.

K. Goebel, P. Bonanni, N. Eklund, “Towards an Integrated Reasoner for Bearings Prognostic”, Proceedings of 2005 IEEE Aerospace Conference, paper z11_0704, 2005.

P. Howard, and J. Reintjes, “A Straw Man for the Integration of Vibration and Oil Debris Technologies”. Helicopter Health and Usage Monitoring Systems Workshop, G. Forsyth, ed., Defense Science and Technology Organisation General Document 197, no.1, pp. 131-136, 1999.

J. Littles and M. Buczek, “Engine System Prognosis”, Proceedings of Materials Science & Technology 2004, September 26-29, 2004, New Orleans, Louisiana, 2004.

A. Loskiewicz-Buczak, R. Uhrig, “Decision Fusion by Fuzzy Set Operations”, Proc. third IEEE Conf. Fuzzy Systems, Vol. 2, pp.1412-1417, 1994.

E. Nadaraya, “On estimating regression”. Theory of Probability and its Applications. 10, 186-190, 1964.

M. Nelson and K. Mason, “A Model-Based Approach to Information Fusion”. Proc. Information, Decision, and Control, pp. 395-400, 1999.

R. Orsagh, J. Sheldon, and C. Klenke, “Prognostics/ Diagnostics for Gas Turbine Engine Bearings”, Proceedings of ASME Turbo Expo 2003, Power for Land, Sea and Air, June 16 – 19, 2003, Atlanta, GA. GT2003-38075.

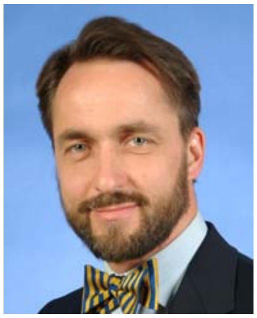
S. Petit-Renaud and T. Denoeux, “Nonparametric Regression Analysis of Uncertain and Imprecise Data Using Belief Functions”, International Journal of Approximate Reasoning, Vol. 35, No. 1, pp. 1-28, 2004.

N. S. V. Rao, “Finite Sample Performance Guarantees of Fusers for Function Estimators, Information Fusion, Vol. 1, no. 1, pp. 35-44, 2000.

P. Smets, “What is Dempster-Shafer’s model?” Advances in the Dempster-Shafer Theory of Evidence, Yager, R., Fedrizzi, M., and Kacprzyk, J., (Eds.), John Wiley & Sons, New York, pp. 5-34, 1994.

G. Watson, "Smooth Regression Analysis", Sankhia, Series A, 26, 359 – 372, 1964.

BIOGRAPHY



***Kai Goebel** received the degree of Diplom-Ingenieur from the Technische Universität München, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively.*

Dr. Goebel joined General Electric's Corporate Research and Development facility in Schenectady, NY in 1997 as a computer scientist after working as a visiting postdoctoral fellow at UC Berkeley from 1996 to 1997. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He has fielded numerous applications for aircraft engines, transportation systems, medical systems, and manufacturing systems. He has published more than 50 technical papers in these areas.

Dr. Goebel is an adjunct professor of the CS Department at Rensselaer Polytechnic Institute (RPI), Troy, NY, since 1998 where he teaches classes in Soft Computing and Applied Intelligent Reasoning Systems.

***Neil Eklund** received B.S. in 1991, two M.S. degrees in 1998, and a Ph. D. in 2002, all at the Rensselaer Polytechnic Institute.*

Dr. Eklund was a research scientist at the Lighting Research Center from 1993 to 1999. He was in the network planning department of PSINet from 1999 to 2002, before joining General Electric Global Research in Niskayuna, NY in 2002. He has worked on a wide variety of research projects, including early detection of cataract using intraocular photoluminescence, multiobjective bond portfolio optimization, and on-wing fault detection and accommodation in gas turbine aircraft engines. His current research interests involve developing hybrid soft/hard computing approaches for real-world problems, particularly real time monitoring, diagnostics, and prognostics.

Dr. Eklund is an adjunct professor in the Engineering/CS department of the Graduate College of Union University, Schenectady, NY, since 2005 where he teaches classes in Computational Intelligence and Machine Learning.

***Pierino Bonanni** received his S.B., S.M., E.E., and Ph.D. degrees in Electrical Engineering from the Massachusetts Institute of Technology in 1983, 1983, 1984, and 1991, respectively.*

Dr. Bonanni joined GE Global Research Center in Niskayuna, New York, in 1991. His specialization is in signal and image processing, with emphasis on detection, estimation, and inverse problems arising in controls and diagnostics applications. He has performed and led a variety of applied research efforts in fault detection. machine vision, robotics, automated manufacturing, satellite navigation and control, non-destructive inspection, and remote sensing. He is the recipient of several GE team and project management awards, and holds 7 U.S. patents.